ABSTRACT OF THE DISCLOSURE

Connected components of dark pixels are clustered from across the image. A "most likely" representative image for each cluster of images is determined, with likelihood determined by a probabilistic model of the image capturing process. An *a*

5   *priori* (prior) probability distributions on bitmaps may be used to determine the most likely representative images. For example, *a priori* probability distributions based on so-called chain codes are implemented. The representative images are used to cluster connected components. Clustering may be repeated. The output page is assembled by replacing each member of a cluster of images by that cluster's representative image.

10